Authors' response

Enactivism, Cognitive Science, and the Jonasian Inference

Dave Ward and Mario Villalobos

1. We are very grateful for this insightful set of commentaries, and thank the authors for taking the time to address some of the issues raised in our target article. Many of the responses are contributions to exactly the kind of conversation within and around the enactive community we hoped our article would start; one that aims at clarifying the complicated set of relationships between the theoretical and methodological commitments of enactivism, phenomenology, and cognitive science. Reading the commentaries together also suggests that, at a couple of crucial stages, we could have presented ideas of our target article more clearly, and we are thankful for the opportunity to try to do that here.

2. What are the points we wanted to make in the target article? Put briefly, we intended to claim that,

- at least since Andreas Weber and Francisco Varela (2002), enactivism has incorporated a theoretical commitment to aspects of Hans Jonas's philosophical biology;
- Jonas's philosophical biology is theoretically committed to an anthropomorphism which is at odds with the methodological commitments of modern science;
- there is thus a tension that needs to be addressed between enactivism's commitment to Jonas's philosophical biology and its aspirations towards becoming a new scientific paradigm.

In the second half of the article we used ideas from Humberto Maturana's autopoietic theory (MAT) to provide an example of an alternative way we might construe the relationship between cognitive science and the lived experience of human and non-human.

3. Though some commentaries (**Maturana** and **Shigeru Taguchi**) raised helpful comments and questions about the part of our target article concerning MAT, this will not be the main focus of our reply. We think we can express the most common themes raised by the commentaries in terms of three questions:

- What is *anthropomorphism*?
- Is it something to which enactivism is really committed?
- Why should it be understood as incompatible with the scientific study of nature?

We will use this reply to say something about each of these crucial questions for our target article in turn.

What is anthropomorphism?

4. Clearly, we left ample room for readers to wonder exactly what *anthropomorphism* refers to in our target article. Is it, **Paulo De Jesus** asks, the attribution of *any* mentality to nonhuman entities (see also **Jean-Luc Petit**'s commentary), or only the attribution of *human-like* mentality to those entities? Are the problems we see for anthropomorphism still problems for a *modern anthropomorphism*, as proposed by **Peter Gaitsch**? Or for the *zoomorphism* in terms of which (as **Steve Torrance** points out) Jonas sometimes characterises his own position?

5. The main objective of our target article is to highlight a tension between the theoretical commitments of enactivism and the methodological commitments of modern science. Given this purpose, what matters for whether a methodology, theory or argument is *anthropomorphic* in our terms is not the *range of properties* attributed to non-human entities, but rather the *way in which our attribution of mental properties is grounded*. The aspect of Jonas's philosophical biology we find problematic is (as **Taguchi** discerns) what we call the *Jonasian Inference* (henceforth "JI"). This is the inference involved when we "take the presence of purposive inwardness in one part of the physical order, viz., in man, as a valid testimony to the nature of that wider reality that lets it emerge" (Jonas 1966: 37). In simple terms, we infer from our first-person experience of teleology that we can take the appearance of teleology in other living organisms at face value.

6. In a chapter of *The Phenomenon of Life* where Jonas is discussing the incompatibility of modern science (Jonas 1966: 72–74) with understanding metabolism as entailing self-perpetuating unities imbued with "needful freedom" (ibid: 80), he asserts that the "[o]rganic identity" of a living organism must be different from the tautologous identity of bits of inert matter with themselves. He asks:

> "But what kind of inference is this? And by whom? How can the unprepared observer infer what no mere analysis of the physical record will ever yield? The unprepared observer cannot […] The observer of life must be prepared by life. In other words, organic existence with its own experience is required of himself for his being able to make that inference […]" (Jonas 1966: 82)

This is the sense of Jonas's claim that life can only be known by life – "happening to be living material things ourselves, we have in our self-experience, as it were, peepholes into the inwardness of [organic] substance […]" (ibid: 91). We might – perhaps should – have written our target article without reference to anthropomorphism, focusing only on the problematic status of JI, and left its central claims intact.

7. In a moment we will say more about the role of JI in contemporary enactivism, and the tension between JI and modern science. Before doing so, let us note that making the role of JI explicit allows us to respond to several of the questions about anthropomorphism raised by the commentators. To **De Jesus**'s question of whether our concern is with the attribution of *human-like* mentality or *any* mentality to non-human entities, we can answer: it depends. JI is most likely to be employed in grounding attributions of human-like mentality, but the questions our target article attempts to raise apply to *any* attribution of mental properties grounded in JI. The compatibility of

the *modern anthropomorphism* proposed by **Gaitsch** with modern science depends on whether such an anthropomorphism avails itself of JI. Finally, **Torrance** is right to note that Jonas was motivated precisely by a concern to avoid *anthropocentrism* in our conception of mind, and thus sometimes speaks of *zoomorphism* instead of anthropomorphism. But such a zoomorphism is equally problematic insofar as its attribution of mental properties is grounded in JI.

Enactivism and the Jonasian Inference

8. In the next section we try to clarify why we find a tension between a commitment to JI and an aspiration to become a new scientific paradigm. But is it actually the case that enactivism grounds its attributions of mentality in JI? **Rasmus Gahrn-Anderson & Matthew Harvey** rightly note that contemporary enactivists argue that the roots of mentality are to be located in dynamic organisational properties such as autopoiesis, autonomy, and adaptivity. Does enactivism not attempt to legitimate its attributions of mental properties via appeal to these features, rather than via JI? **Jérôme Proulx** finds no such commitment in his own enactivist work on mathematical cognition, nor in the related work of his colleagues.

9. However, contra **Gahrn-Anderson & Harvey**, we think that the commitment to a deep continuity between life and mind that is characteristic of much contemporary enactivism is more often motivated by JI than by a belief that reductive explanation of teleological phenomena in terms of autopoietic or adaptive dynamics has been provided. To repeat a quote from our target article, consider Weber and Varela's claim that:

> "[B]efore being scientists we are first living beings, and as such *we have evidence of our intrinsic teleology in us*. And, in observing other creatures struggling to continue their existence – starting from simple bacteria that actively swim away from a chemical repellent – we can, *by our own evidence*, *understand teleology as the governing force of the realm of the living*." (Weber & Varela 2002: 110, emphasis added)

Here, and in the rest of Weber and Varela's paper, a teleological understanding of living systems is grounded in the evidence of our own experience, via JI, and it is only in light of this phenomenological evidence that autopoietic or adaptive dynamics present themselves as plausible dynamical underpinnings of teleological properties. Similarly, Evan Thompson holds that:

> "[T]he theory of autopoiesis provides a naturalistic interpretation of the teleological conception of life originating in experience, but our experience of our own bodily being is a condition of possibility for our comprehension of autopoietic selfhood." (Thompson 2007: 164, emphasis added)

And in Ezequiel Di Paolo's influential paper "Autopoiesis, adaptivity, teleology, agency" he departs from Weber and Varela (2002) in arguing that enactivists should ground teleology in adaptive rather than autopoietic dynamics, while sharing their commitment to JI:

> "[T]he attribution of teleology to metabolism is justified partly by means of intuition outside scientific discourse. Jonas implicitly admits that establishing in metabolism the breaking point between extended neutral processes and concernful identity is a matter of appropriate choice.

> How are we to justify this and further choices, or question their sufficiency, when the criteria of validation are, at least partly, outside science? The answer must be: by the use of phenomenological insight or other disciplined intuitions." (Di Paolo 2005: 431f)

10. Each of these quotes make clear that, at least for these canonical enactivist thinkers, JI is called upon to do important work *before* autopoietic or adaptive dynamics can appear as candidate explanations of teleological properties. Might enactivists dispense with JI and attempt a straightforward reductive explanation of teleological properties in terms of dynamical organisation? This question deserves more discussion than we can provide here. But a common theme in recent work on enactivism is the provision of arguments that autopoietic or adaptive dynamics are ill-suited to this explanatory role (see, e.g., Villalobos & Ward 2015; De Jesus 2016; Barrett 2015; Barandiaran 2016).

11. Philosophers of mind and cognition have spent much of the past five decades attempting to explain teleological properties of our mental states in terms of structural, functional or dynamical properties. It is fair to say that the consensus is that these attempts have yet to succeed, and it is not clear how adaptive or autopoietic accounts can help with the problems they have faced – for example, how can an account in terms of structure, function or dynamics adequately specify success conditions for teleological states? How can such an account accommodate the normativity that separates the genuinely teleological character of some mental states from a mere covariation relation between a state of an organism and a state in the world? It may be that the conceptual apparatus already at the disposal of enactivists can yield satisfactory responses to questions such as these – but further work is required to show this. For the enactivist authors quoted above, however, these questions are misplaced. None of these authors aim at the reductive explanatory goal of showing how teleology emerges from non-teleological properties of structure, function or dynamics. Instead, via JI, they argue that we should understand particular forms of dynamical organisation as imbued with immanent teleology. We think it is fair to say that these authors have set the research agenda for most contemporary enactivist work. Might enactivism nonetheless dispense with JI? We will return to this question below. But now that we have clarified our conception of JI and the use enactivism makes of it, we can say more about its problematic relationship with modern science.

Cognitive science and the Jonasian Inference

12. Several commentators (**De Jesus, Gaitsch, Petit, Torrance**) raise questions about why employing JI should be understood as incompatible with the scientific study of nature, and others (**Maturana, Taguchi**) suggest specific ways in which the relationship between phenomenology and science might be understood which go beyond the discussion of our target article. Hopefully the above clarifications about the role of JI in our argument and in enactivism already suggest responses to some of these questions. For example, in response to **Petit** (and to parts of the commentaries of **De Jesus** and **Torrance**) it should now be clear that our concern in our target article is not with the attribution of teleological properties to nonhuman organisms *tout court*, but with grounding that attribution in JI.

13. So, just what is the incompatibility we see between the methodological commitments of *modern science* and a commitment to JI? In our target article we mainly rely on pointing towards Jonas's own case for this incompatibility – one reason for this is to encourage enactivists who make use of Jonas's work to address this issue. If we take Jonas's work seriously enough to absorb a commitment to JI from it, then they should also engage with the parts of Jonas's work dealing with the tension between endorsing immanent teleology on the basis of JI and modern science. Another reason for our relying on Jonas for this purpose is simply that specifying the methodological commitments of modern science is a difficult task – perhaps an impossible one, due to the heterogeneity of the research programmes we are willing to classify as scientific. One response to this difficulty is to meet it head on, by engaging in the kind of reflection on scientific method found in **Maturana**'s interesting commentary. However, we do not think that the case we present in our target article requires us to do this.

14. When we speak, following Jonas, of *modern* science the intended contrast is with a *pre-modern*, Aristotelian conception of science, which conceives of the natural world as populated with irreducible intrinsic teleology. As we noted, an Aristotelian explanation of why rocks fall through air while fire rises proceeds by attributing a natural *telos* to these entities, so that rocks *strive* to be close to the earth while fire strives for the heavens. We also noted that the unacceptability of such an explanation by the standards of modern science is not due to the choice of particular elements in this theory, or with the particular *telos* that characterizes each of them. Rather, it is with attributing *irreducible* natural purposes to entities in our explanation; that is, purposes that receive no further explanation in terms of structure, function or dynamics. Above, we left the question of whether enactivism might, in future, provide such a further explanation open. But we also claimed that enactivists like Varela (Weber & Varela 2001), Di Paolo (2005) and Evan Thompson (2007) do not see the provision of such explanations as their task. When autopoietic or adaptive dynamics are identified with immanent purposiveness by such enactivists, this is not because an explanation of the emergence of natural purposes from non-teleological dynamics has been given. Rather, JI has been employed to allow us to understand the relevant dynamics as already imbued with teleology.

15. We find it difficult to add much more to this by way of explaining the incompatibility between modern science and the acceptance of intrinsic natural purpose, or immanent teleology, that is the result of JI. This is because we agree with Jonas's assessment that this rejection is not an empirical result obtained by science, but rather a methodological presupposition that demarcates the boundaries of scientific inquiry. This is why mainstream philosophy of cognitive science has been preoccupied for the last fifty years by trying (and failing) to provide reductive explanations of teleological properties of mental states in terms of structural and functional properties. Examples include Fred Dretske's (e.g., 1981) information-theoretic theory of content and Ruth Millikan's (e.g., 1984) teleosemantics. A genuinely naturalistic explanation of teleological phenomena, such thinkers believe, must explain them in terms of non-teleological states, structurally and functionally construed. Above, we have tried to clarify that enactivism's endorsement of JI absolves enactivism of the responsibility for

providing such further explanations, and it is this that puts it in tension with modern scientific method.

16. However, is this not to ignore the very fact that our target article is supposed to concern: that enactivism frequently presents itself as a *new paradigm* for cognitive science? As **Taguchi** and **Torrance** rightly point out, a belief in the need for the reciprocal circulation of ideas between phenomenology and cognitive science is one of the founding principles of enactivism. So is the conception of cognitive science and its methodology just sketched, where we aspire to reductively explain every phenomenological property or structure in terms of structure, function or dynamics, not just what enactivism has always sought to rid us of? We agree that we did not say enough about enactivism's commitment to a *reciprocal* interplay between phenomenology and cognitive science in our original target article. Similarly, when we presented the quotes from Weber and Varela, Di Paolo and Thompson above, we did not mention that each of them occurs in the context of the authors stressing the need for a *dialogue* between phenomenology and cognitive science. The way we have presented things in our target article and in this response has so far been in terms of an opposition between science, which begins with structure, function, and dynamics and attempts to explain further properties and phenomena in these terms, and phenomenology, which begins with our lived experience and invites us to view the material world in these terms. But is the appeal of enactivism as a paradigm not supposed to lie in its showing us a middle way between these two extremes?

Commented [a1]: redundant?

17. Once again, we cannot attempt to fully address this important issue here (nor do we know how to do so). But we think that considering the nature of the interplay between phenomenology and cognitive science, and the delicate task of reconciling the methodological commitments of each, is where the enactive community should focus its efforts. Here is one way we can restate the central claim of our target article in light of this: contemporary enactivism's conception of the relationship between phenomenology and cognitive science is not sufficiently reciprocal. As we see it, phenomenology is calling the shots in this relationship in a problematic way. In endorsing JI, contemporary enactivism begins with our own experience, then demands that the structural, functional and dynamical properties that are the explanatory materials of cognitive science be reconstrued in its terms. Consider an alternative relationship, where science calls the shots. Science would begin with non-sentient and non-teleological structures and processes, then demand that the subject matter of phenomenology be reconceived in these terms. Instead of taking our experience of teleology at face value and reconceiving the material world in light of this, we would be guided by science's non-sentient, non-teleological conception of material reality, and conclude that our experience of teleology and subjectivity is illusory. This conception of the priority of scientific understanding is incompatible with a genuinely *reciprocal* relationship between cognitive science and phenomenology. Likewise, we think that in relying on JI, contemporary enactivists commit themselves to a conception of the priority of phenomenological understanding that is incompatible with the reciprocal relationship between phenomenology and cognitive science at which enactivism originally aimed.

18. There are many other important, insightful and interesting points in the commentaries on our target article that we have not been able to address here. We apologise for this, and hope to take up some of these issues in future work. Nonetheless, we hope this response has served to clarify the most important points we wanted to make in our target article, and that it will motivate some readers to take up the challenges we have tried to pose for enactivism.

References

Barandiaran X. E. (2016) Autonomy and enactivism. Towards a theory of sensorimotor autonomous agency. Topoi. In press. http://cepa.info/2471

Barrett N. F. (2015) The normative turn in enactive theory: An examination of its roots and implications. Topoi. Online first. http://cepa.info/2473

De Jesus P. (2016) Autopoietic enactivism, phenomenology and the deep continuity between life and mind. Phenomenology and the Cognitive Sciences. In press. http://cepa.info/2385

Di Paolo E. A. (2005) Autopoiesis, adaptivity, teleology, agency. Phenomenology and the Cognitive Sciences 4(4): 429–452. http://cepa.info/2269.

Dretske F. (1981) Knowledge and the flow of information. MIT Press, Cambridge MA.

Jonas H. (1966) The phenomenon of life. Harper & Row, New York.

Millikan R. G. (1984) Language, thought and other biological categories. MIT Press, Cambridge MA.

Thompson E. (2007) Mind in life: Biology, phenomenology, and the sciences of mind. Harvard University Press, Cambridge MA.

Villalobos M. & Ward D. (2015) Living systems: Autopoiesis, autonomy and enaction. Philosophy & Technology 28(2): 225–239. http://cepa.info/2511

Weber A. & Varela F. (2002) Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. Phenomenology and the Cognitive Sciences 1: 97–125. http://cepa.info/2087

Received: 29 February 2016

Accepted: 6 March 2016